

Fixed-delay Events in Generalized Semi-Markov Processes Revisited

In the area of performance evaluation, the performance of real-life systems such as queues, assembly lines or communication protocols are analyzed, using formal mathematical models, such as Generalized Semi-Markov Processes (GSMP). We consider a previously studied class of GSMP extended with events that occur after a fixed delay. Disproving several previous results, we show that such GSMP can exhibit a surprising(ly?) unstable behavior in the long-run. The instability is caused by properties of the model not found in real-life. To avoid this undesirable situation, we also provide syntactical conditions upon which a GSMP model has always a stable behavior.

Further reading:

http://en.wikipedia.org/wiki/Queueing_theory - Related area explaining one application of the theory.

http://en.wikipedia.org/wiki/Performance_engineering - Loosely related, performance aspects from the point of view of software engineers. In fact, there is not much about performance modelling in the article.

Creating corpora in few clicks using Corpus Architect

This technical report describes Corpus Architect -- a web application for building textual corpora with a minimal effort. The application loads data into corpus manager Manatee/Bonito[1], thus enables language researchers to use the powerful corpus manager with their own data.

Corpus Architect **builds** corpora from input documents in various text formats. It **can also gather** data from the web and apply advanced boilerplate removal and deduplication algorithms. In this scenario, an approach similar to Corpus Factory[2] is used: the user supplies just a few keywords, Corpus Architect uses an external web search engine to look up documents and builds the corpus. Postprocessing tools such as tokenization, lemmatization and part of speech tagging are available for English, German, French, Spanish, Chinese, Japanese and other languages.

Corpus Architect was used by 500 users to build 1100 corpora containing 23000 files of total size reaching 50 GB in 2012.

???final?

Time Features in SCStudio

Message Sequence Charts (MSC) are used in design phase as an intuitive formalism to describe distributed systems. The Sequence Chart Studio is an open source user friendly drawing and verification tool for MSC. SCStudio supports specification and verification of MSC with time concepts. We define the syntax and semantics of timed MSC and explain algorithms which work with them, i.e. time consistency, tightening and time race checker.

???final??

PAN 2012 - competition on plagiarism detection

Plagiarism has become serious problem mainly because of the electronically available documents. This situation has given rise to many anti-plagiarism tools. The PAN is an international competition on plagiarism detection. In this talk, I will introduce the settings of PAN 2012 competition. We competed in both the candidate document retrieval and the detailed comparison tasks. In this competition, we used methodology based on three-way search engine queries with multi-feature document comparison. Our candidate retrieval system was based on extraction of three different types of Web queries (keywords-based; intrinsic plagiarism based; and headers based) with narrowing their execution by skipping certain passages of an input document. The detailed comparison approach led to the second place in the detailed document comparison subtask. In the candidate document retrieval subtask our approach won in majority of the performance measures. We discovered that keywords based queries are the most promising and the headers based queries do not pay off in the overall performance.

Complex event processing in building management systems

Complex event processing (CEP) is relatively new approach to real-time monitoring and analysis of large systems. Nowadays, CEP is used mostly in fields of network security, fraud detection, algorithmic stock trading and business process monitoring. This talk focuses on use of CEP technologies in the field of Facility management, namely in the area of Building management systems. It describes typical use of CEP in BMS and ways to process building management data. Existing implementation of such system is briefly presented and capabilities of the system on "real life" use cases are illustrated.

http://en.wikipedia.org/wiki/Complex_event_processing

Designing onboarding on mobile applications

Onboarding is a process of helping people get started with a product or service and turning them from one-time visitors to returning customers during the first interaction. Designing onboarding on mobile applications presents unique challenges, because of the short attention span of mobile users, high bounce rate of mobile applications and low discoverability of touch interactions. In this presentation, I will talk about different strategies of onboarding on mobile applications, their advantages and disadvantages and best practices for design.